

Desvio Padrão - Análise de Acidentes Em Minas Gerais

Análise Descritiva da Distribuição de
Acidentes entre as Rodovias do Estado

Arthur Damasceno Dalvino

Desafio Projeto Simulado - EstatMG
Quarto, Mapas & Tabelas Interativos
Análise Temporal, Rodovias Mineiras

Brasil, Setembro de 2025



Desvio Padrão - Análise de Acidentes Em Minas Gerais

Análise Descritiva da Distribuição de Acidentes entre as Rodovias do Estado

Arthur Damasceno Dalvino

Orientadora: Isadora Silva Drumond
Diretora do Setor de Projetos na EstatMG.

Desafio Projeto Simulado - EstatMG
Quarto, Mapas & Tabelas Interativos
Análise Temporal, Rodovias Mineiras

Projeto Desenvolvido para o Processo Seletivo da Empresa Júnior

Brasil, Setembro de 2025

Desvio Padrão - Análise de Acidentes Em Minas Gerais

Copyright © 2025 - Arthur Damasceno Dalvino, Desafio Projeto Simulado - EstatMG.

Este projeto é um trabalho original, elaborado exclusivamente para este fim, e todos os autores cujos estudos e publicações contribuíram para sua realização foram devidamente citados. Os direitos sobre este trabalho pertencem à *EstatMG*.



Preparation of this work was facilitated by the use of the *IPLeiria-Thesis* template.

Resumo

Com base na análise de acidentes rodoviários de 2020, os principais pontos de risco foram identificados, revelando padrões geográficos, temporais e de condições específicas.

A análise mostra que, embora Betim tenha o maior volume de acidentes, Patos de Minas apresenta a maior taxa de mortalidade. As rodovias BR-381 e BR-040 são as mais perigosas, com quilômetros específicos na BR-040 apontados como de alta criticidade. O tipo de acidente mais comum é a “saída de leito carroçável”.

Fatores como o clima e o tipo de pista também são determinantes: a neblina e as pistas simples estão associadas a uma maior severidade dos acidentes.

Temporalmente, os finais de semana concentram a maior frequência de ocorrências. O horário mais letal é a madrugada, apesar de a maioria dos acidentes ocorrer nos horários de pico.

Observou-se também uma queda atípica em abril, possivelmente pela pandemia de COVID-19, e um pico em dezembro, relacionado às viagens de fim de ano.

Declaração sobre o Uso de Inteligência Artificial

i Exemplo de Utilização

Reconheço a utilização das ferramentas de inteligência artificial Gemini (<https://gemini.google.com>) e Claude AI (<https://claude.ai>) para melhorar a precisão linguística, a organização do documento e para auxiliar na criação e estilização de tabelas e gráficos neste trabalho.

Descrição da Utilização

Instrução 1: Analise o Texto, corrija a gramática e ajuste a coesão, sem alterar a forma majoritária do texto.

Instrução 2: Transforme esta Tabela em uma tabela LaTeX seguindo este modelo.

Instrução 3: Quero ajustar estes gráficos Ggplot e estas tabelas Reactable para ficarem mais atrativos visualmente, como posso realizar estas mudanças?

Conteúdo

<i>Lista de Figuras</i>	iv
<i>Lista de Tabelas</i>	1
1 Entendimento do Dados e Contexto	2
1.1 Análise Preliminar dos Dados	2
1.1.1 Estatísticas Descritivas	3
1.2 Planejamento da Análise	5
2 Análise Exploratória de Dados (EDA)	6
2.1 Análise Temporal dos Acidentes	6
2.1.1 Distribuição Semanal, Mensal e Horária	6
2.2 Análise das Causas e Tipos de Acidente	9
2.2.1 Condições da Via e do Clima	9
2.2.2 Tipos de Acidente Mais Frequentes e Letais	9
2.3 Análise de Severidade e Geográfica	12
2.3.1 Métricas Gerais e Severidade	12
2.3.2 Pontos Críticos: Municípios e Rodovias	14
3 CONclusão	16

Lista de Figuras

1.1	Matriz de Correlação entre as variáveis numéricas	4
1.2	Matriz de Associação entre as variáveis Categóricas	4
2.1	Média diária de acidentes por dia da semana.	7
2.2	Evolução do total de acidentes por mês em 2020.	8
2.3	Heatmap da concentração de acidentes por dia da semana e fase do dia. . .	9
2.4	Frequência de acidentes (barras) versus taxa de mortalidade (linha) por condição climática.	10
2.5	Heatmap da frequência de acidentes por tipo e pista.	11
2.6	Top 10 tipos de acidente mais letais.	12
2.7	Distribuição do número de mortos em acidentes fatais.	13
2.8	Frequência de acidentes (barras) versus taxa de mortalidade (linha) por hora.	14

Lista de Tabelas

1.1	Resumo detalhado das variáveis do dataset de acidentes.	2
1.2	Casos de acidentes com informação da BR faltante.	3
2.1	Total e Média Diária de Acidentes por Dia da Semana (2020).	6
2.2	Análise Percentual e Acumulada de Acidentes por Mês (2020).	7
2.3	Total de Acidentes por Hora do Dia (2020).	8
2.4	Severidade dos Acidentes por Condição Meteorológica (2020).	9
2.5	Severidade dos Acidentes por Tipo de Pista (2020).	10
2.6	Distribuição de Frequência por Tipo de Acidente (2020).	10
2.7	Taxa de Mortalidade por Tipo de Acidente (2020).	11
2.8	Métricas Chave de Severidade dos Acidentes (2020).	12
2.9	Totais Gerais de Vítimas em Acidentes (2020).	12
2.10	Amostra da Taxa de Mortalidade por Hora do Dia (2020).	13
2.11	Top 10 Municípios com Maior Número de Mortos (2020).	14
2.12	Top 10 Rodovias Federais (BRs) por Total de Ocorrências (2020).	15
2.13	Top 10 Pontos Críticos (BR-KM) por Número de Ocorrências (2020).	15

1

Entendimento do Dados e Contexto

Official Repository: *GitHub Repository*

Este relatório apresenta uma análise aprofundada sobre as ocorrências de acidentes nas rodovias de Minas Gerais.

1.1 Análise Preliminar dos Dados

Os dados já estão bem organizados e é possível inferir facilmente o significado das colunas. Um possível ponto de confusão é o formato da data (`data_inversa`), que segue o padrão internacional (ano/mês/dia).

Tabela 1.1: *Resumo detalhado das variáveis do dataset de acidentes.*

Informação	Váriavel
Onde e Quando	id data dia horário uf br (6% NA) km (6% NA) município
Como Aconteceu	tipo_acidente fase_dia condicao_metereologica tipo_pista
Impacto do Acidente	pessoas mortos feridos ileso veiculos
Localização Precisa	latitude longitude

Além disso, existem pouquíssimos valores ausentes (NA), concentrados nas colunas `br` e `km`, que se referem, respectivamente, à rodovia e ao quilômetro onde o acidente ocorreu. Essas ausências correspondem a cerca de 6,12% do total de observações. Tais observações não aparentam ter um padrão em comum, seja em localização, número de vítimas ou tipo de acidente, o que sugere que podem ser falhas no momento do registro dos dados ou acidentes em locais não formalmente catalogados.

Tabela 1.2: Casos de acidentes com informação da BR faltante.

Data	Dia da Semana	Horário	UF	Município	Tipo de Acidente
2020-01-01	quarta	15:55	MG	SABARA	Tombamento
2020-05-29	sexta	16:15	MG	JUIZ DE FORA	Colisão traseira
2020-06-12	sexta	12:00	MG	JOAO MONLEVADE	Capotamento
2020-07-10	sexta	07:41	MG	CALDAS	Colisão transversal

1.1.1 Estatísticas Descritivas

A análise das estatísticas descritivas das variáveis numéricas não revelou, à primeira vista, dados suspeitos ou a existência de outliers evidentes, considerando valores como o número máximo e mínimo de feridos, o desvio padrão e a variância. Alguns achados interessantes envolvem os veículos, uma vez que a mediana é 1,0, enquanto a média é 1,55. Isso significa que mais de 50% dos acidentes registrados envolvem apenas um único veículo.

O número máximo de pessoas envolvidas é de 14; no entanto, a média é de 2,23 pessoas. Considerando as mortes e os feridos graves, a mediana, o primeiro e o terceiro quartil são zero. Portanto, pelo menos 75% de todos os acidentes neste dataset não tiveram vítimas fatais.

A matriz de correlação (Método de Spearman) indicou, em sua maioria, relações consideradas óbvias entre as variáveis. Como no correlogramas abaixo

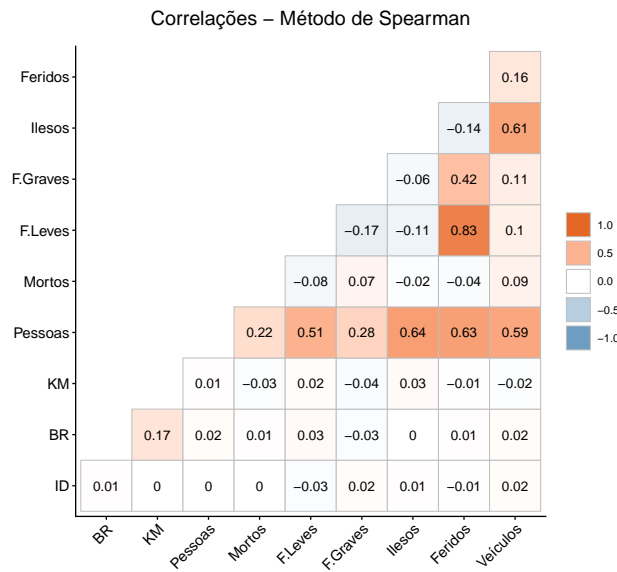


Figura 1.1: Matriz de Correlação entre as variáveis numéricas

As correlações mais fortes estão relacionadas ao número de pessoas envolvidas e suas consequências imediatas. Logicamente, se há mais pessoas, haverá mais veículos, feridos e ilesos. O restante das correlações observadas parece ser bem razoável.

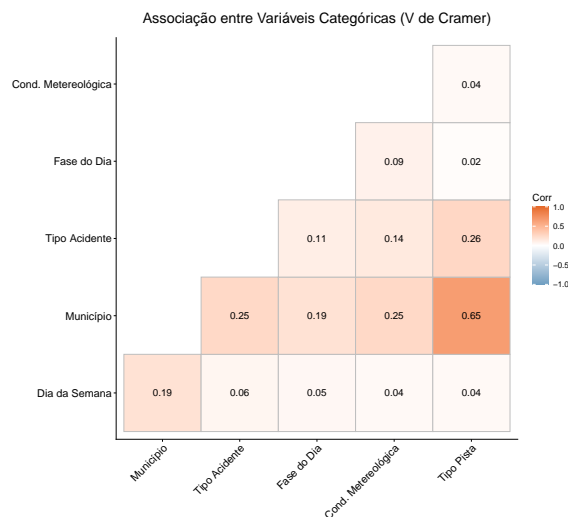


Figura 1.2: Matriz de Associação entre as variáveis Categóricas

As medidas de associação, por outro lado, trazem resultados mais curiosos. O município onde aconteceu o acidente parece ter relação com a maioria das variáveis, o que pode significar que municípios com estruturas viárias piores ou mais antigas influenciam o tipo de ocorrência. Além disso, o tipo de pista parece estar bem associado ao tipo de acidente.

1.2 Planejamento da Análise

Diante desse cenário e para realizar tanto as análises solicitadas quanto outras explorações pertinentes, elaboramos os seguintes eixos para guiar o trabalho.

Objetivo do Cliente

“Meu objetivo é avaliar as principais causas de acidentes nas rodovias de Minas Gerais, bem como um mapeamento dos tipos de acidentes e das rodovias. Também gostaria de entender os acidentes de acordo com variáveis relativas a tempo: dia da semana e mês.”

Eixos de Análise Propostos

1. Análise das Causas de Acidentes:

- (a) Distribuição por condição meteorológica.
- (b) Distribuição por tipo de pista.
- (c) Distribuição por tipo de acidente

2. Análise de Variáveis Temporais:

- (a) Evolução mensal do número de acidentes.
- (b) Distribuição de acidentes por dia da semana.
- (c) Distribuição por faixa de horário (manhã, tarde, noite, madrugada).

3. Análise de Variáveis Geográficas:

- (a) Distribuição geográfica de mortos e feridos.
- (b) Ranking das 10 principais rodovias (BRs) com maior número de ocorrências.

4. Métricas e Proporções Chave:

- (a) Taxa de mortalidade
- (b) Taxa de feridos graves
- (c) Taxa de ileso

2

Análise Exploratória dos Dados (EDA)

Neste capítulo, mergulhamos na análise exploratória dos dados de acidentes rodoviários de 2020. O objetivo é dissecar os padrões temporais, geográficos e as causas subjacentes às ocorrências, seguindo uma estrutura lógica que parte da análise da qualidade dos dados até as métricas de severidade e os pontos críticos nas rodovias.

2.1 Análise Temporal dos Acidentes

A dimensão temporal é crucial para entender quando os acidentes são mais prováveis de ocorrer.

2.1.1 Distribuição Semanal, Mensal e Horária

A Tabela 2.1 e a Figura 2.1 revelam que os acidentes se concentram nos fins de semana, com picos na **sexta-feira** e no **sábado**, dias de maior fluxo de veículos para lazer e viagens.

Tabela 2.1: Total e Média Diária de Acidentes por Dia da Semana (2020).

Dia da Semana	Total de Acidentes	Média Diária
domingo	1053	20.2
segunda	857	16.5
terça	794	15.3
quarta	825	15.6
quinta	846	16.0
sexta	1059	20.4
sábado	1102	21.2

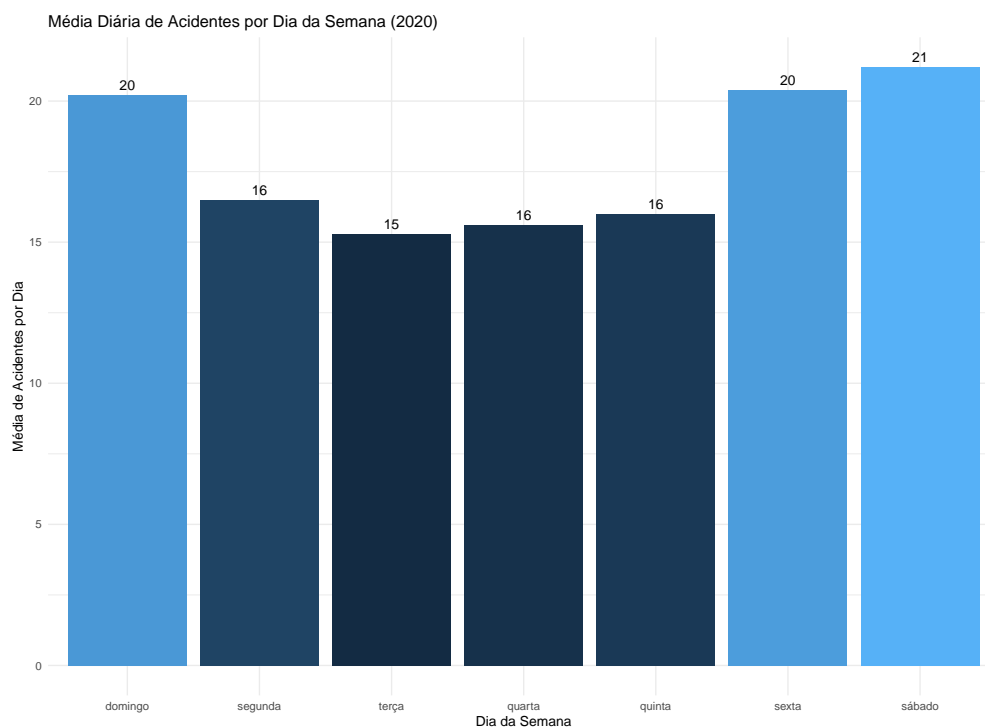


Figura 2.1: Média diária de acidentes por dia da semana.

A análise mensal, detalhada nas Tabelas e 2.2 e visualizada na Figura 2.2, mostra um pico em **dezembro** (10.2% do total anual) e uma queda atípica em **abril**, reflexo da pandemia de COVID-19.

Tabela 2.2: Análise Percentual e Acumulada de Acidentes por Mês (2020).

Mês	Nº de Acidentes	% do Total no Mês	% Acumulado no Ano
janeiro	619	9.47	9.47
fevereiro	561	8.58	18.10
março	450	6.88	24.90
abril	378	5.78	30.70
maio	475	7.27	38.00
junho	480	7.34	45.30
julho	496	7.59	52.90
agosto	580	8.87	61.80
setembro	630	9.64	71.40
outubro	645	9.87	81.30
novembro	554	8.48	89.80
dezembro	668	10.20	100.00

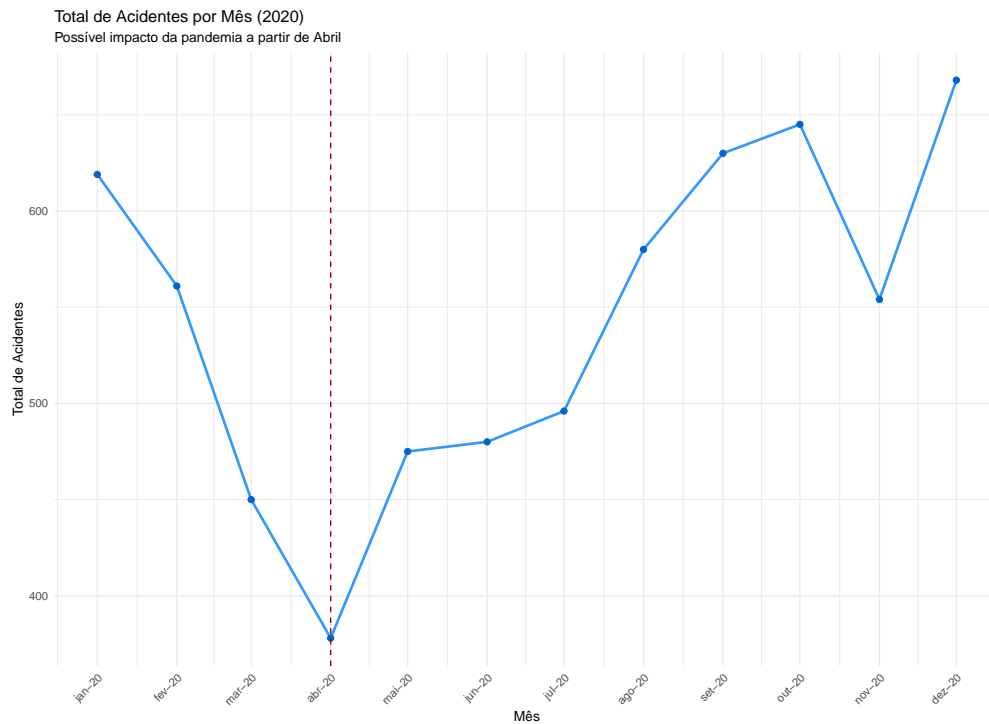


Figura 2.2: Evolução do total de acidentes por mês em 2020.

A distribuição horária (Tabela 2.3) e o cruzamento com a fase do dia (Figura 2.3) indicam que os horários de pico de tráfego, como o final da tarde, concentram o maior número de ocorrências.

Tabela 2.3: Total de Acidentes por Hora do Dia (2020).

Hora	Total	Hora	Total
0	133	12	312
1	111	13	320
2	106	14	343
3	97	15	361
4	125	16	406
5	180	17	487
6	306	18	511
7	340	19	409
8	291	20	292
9	281	21	240
10	288	22	214
11	329	23	187

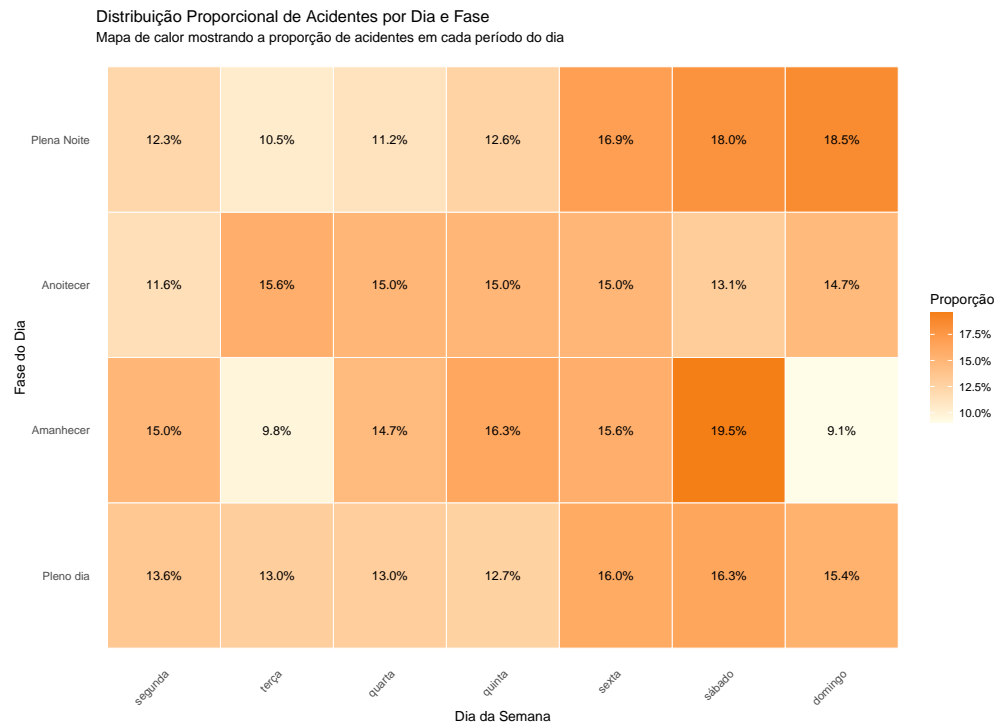


Figura 2.3: Heatmap da concentração de acidentes por dia da semana e fase do dia.

2.2 Análise das Causas e Tipos de Acidente

Compreender os fatores contribuintes é essencial para a prevenção.

2.2.1 Condições da Via e do Clima

A Tabela 2.4 e a Figura 2.4 mostram que, embora a maioria dos acidentes ocorra em **Céu Claro**, a taxa de mortalidade é drasticamente superior sob **Nevoeiro/Neblina** (7.62%), evidenciando o alto risco dessa condição.

Tabela 2.4: Severidade dos Acidentes por Condição Meteorológica (2020).

Condição	Frequência	Mortos	Pessoas	Mortos/Ac.	Mortalidade (%)
Céu Claro	3960	329	8676	0.083	3.79
Nublado	1225	81	2676	0.066	3.03
Chuva	1208	96	2907	0.079	3.30
Ignorado	100	7	209	0.070	3.35
Nevoeiro/Neblina	43	8	105	0.186	7.62

Da mesma forma, a Tabela 2.5 indica que pistas do tipo **Simplex** concentram a maior taxa de mortalidade (4.92%), mais que o dobro da registrada em pistas do tipo **Dupla**.

2.2.2 Tipos de Acidente Mais Frequentes e Letais

A Tabela 2.6 lista a frequência de cada tipo de acidente, com a **Saída de leito carroçável** sendo a mais comum (25.7%). A Figura 2.5 cruza essa informação com o tipo de pista.

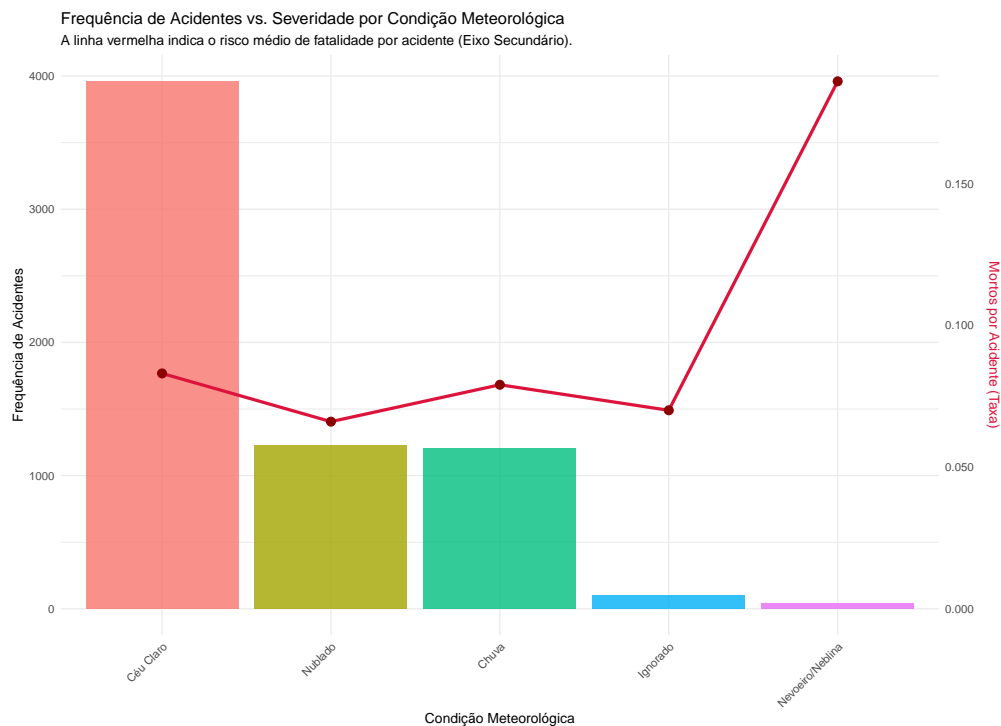


Figura 2.4: Frequência de acidentes (barras) versus taxa de mortalidade (linha) por condição climática.

Tabela 2.5: Severidade dos Acidentes por Tipo de Pista (2020).

Tipo de Pista	Frequência	Mortos	Pessoas	Mortos/Ac.	Mortalidade (%)
Simples	3335	383	7784	0.115	4.92
Dupla	2990	132	6342	0.044	2.08
Múltipla	211	6	447	0.028	1.34

Tabela 2.6: Distribuição de Frequência por Tipo de Acidente (2020).

Tipo de Acidente	Frequência	Percentual (%)
Saída de leito carroçável	1682	25.70
Colisão traseira	1066	16.30
Tombamento	869	13.30
Colisão lateral	622	9.52
Colisão transversal	597	9.13
Colisão frontal	543	8.31
Queda de ocupante de veículo	312	4.77
Capotamento	264	4.04
Atropelamento de Pedestre	223	3.41
Incêndio	150	2.29
Atropelamento de Animal	125	1.91
Engavetamento	83	1.27

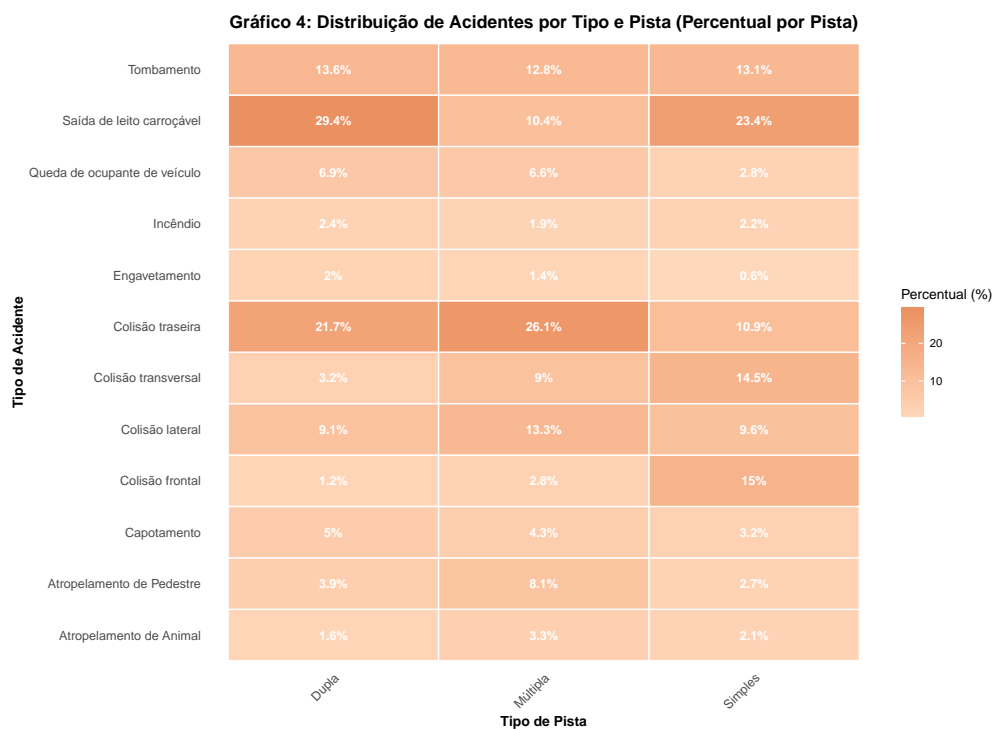


Figura 2.5: Heatmap da frequência de acidentes por tipo e pista.

Contudo, a Tabela 2.7 e a Figura 2.6 mostram que a letalidade conta uma história diferente. **Colisão frontal** e **Atropelamento de Pedestre** são os tipos de acidente mais letais, ambos com uma taxa de mortalidade de 10.8%.

Tabela 2.7: Taxa de Mortalidade por Tipo de Acidente (2020).

Tipo de Acidente	Total Pessoas	Total Mortos	Taxa Mortalidade (%)
Atropelamento de Pedestre	498	54	10.80
Colisão frontal	1820	196	10.80
Tombamento	1188	49	4.12
Saída de leito carroçável	3029	89	2.94
Atropelamento de Animal	234	6	2.56
Capotamento	522	12	2.30
Colisão transversal	1608	36	2.24
Colisão traseira	2921	49	1.68
Queda de ocupante de veículo	428	6	1.40
Colisão lateral	1678	21	1.25

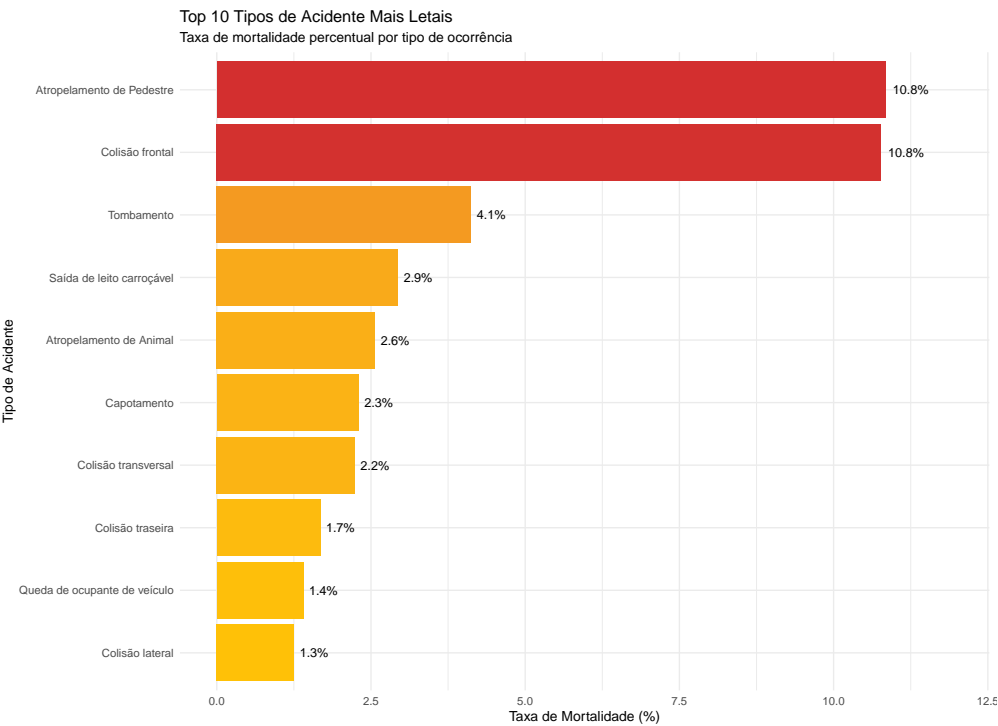


Figura 2.6: Top 10 tipos de acidente mais letais.

2.3 Análise de Severidade e Geográfica

2.3.1 Métricas Gerais e Severidade

As métricas gerais (Tabelas 2.8 e 2.9) mostram uma taxa de mortalidade geral de 3.58%. A Figura 2.7 revela que a maioria dos acidentes fatais resulta em uma única vítima.

Tabela 2.8: Métricas Chave de Severidade dos Acidentes (2020).

Métrica	Valor (%)	Total Absoluto
Taxa de Mortalidade	3.58	521
Taxa de Feridos Graves	12.66	1845
Taxa de Ilesos	39.70	5785

Tabela 2.9: Totais Gerais de Vítimas em Acidentes (2020).

Total Pessoas	Total Mortos	Total Feridos Graves	Total Ilesos
14573	521	1845	5785

A Tabela 2.10 e a Figura 2.8 destacam um padrão crucial: embora a frequência de acidentes seja maior no final da tarde, a **madrugada** é o período com a maior taxa de mortalidade, atingindo um pico de 12.4% à 1h da manhã.

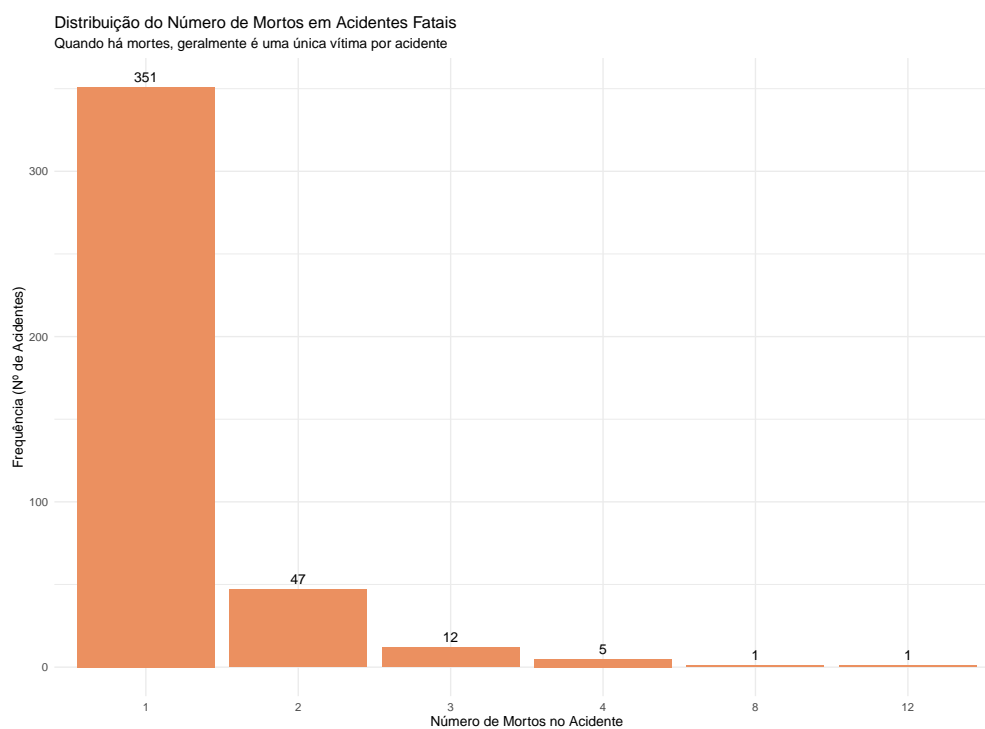


Figura 2.7: Distribuição do número de mortos em acidentes fatais.

Tabela 2.10: Amostra da Taxa de Mortalidade por Hora do Dia (2020).

Hora do Dia	Total Pessoas	Total Mortos	Taxa de Mortalidade (%)
0	289	19	6.57
1	226	28	12.40
2	207	11	5.31
3	200	11	5.50
4	245	20	8.16
5	361	20	5.54
6	703	19	2.70
7	789	18	2.28
8	650	8	1.23
9	607	12	1.98

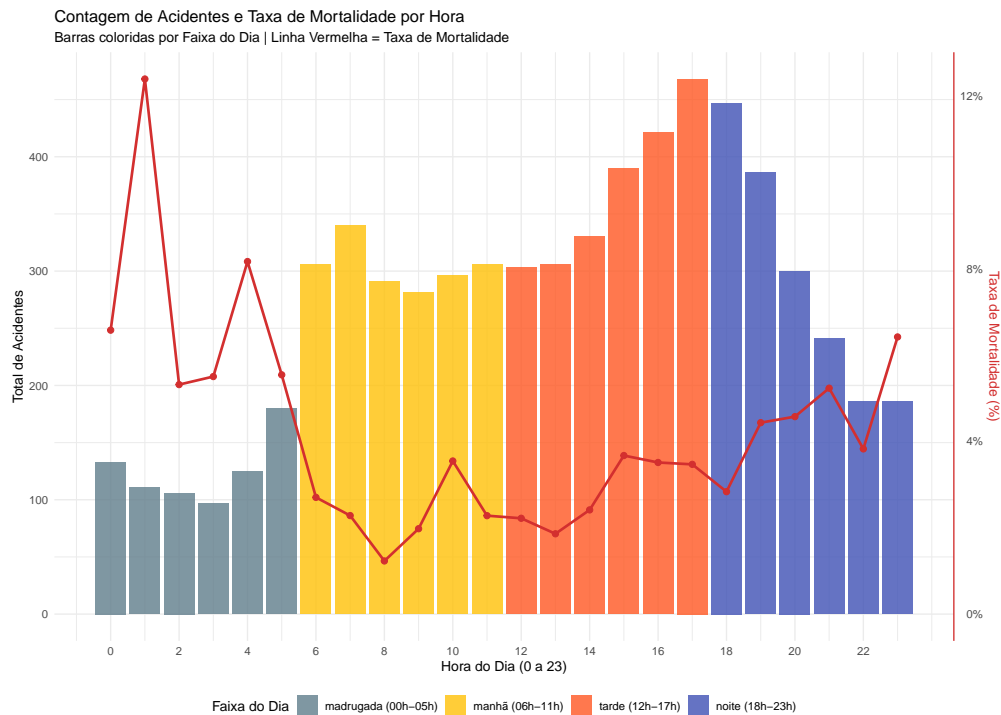


Figura 2.8: Frequência de acidentes (barras) versus taxa de mortalidade (linha) por hora.

2.3.2 Pontos Críticos: Municípios e Rodovias

Geograficamente, a Tabela 2.11 mostra que **Patos de Minas** lidera em número de mortos. As Tabelas 2.12 e 2.13 identificam as rodovias **BR-381** e **BR-040** como as mais perigosas, com destaque para os quilômetros 510 e 511 da BR-040, que sozinhos somam 33 ocorrências.

Tabela 2.11: Top 10 Municípios com Maior Número de Mortos (2020).

Município	Mortos	Feridos	Acidentes
PATOS DE MINAS	17	63	51
GOVERNADOR VALADARES	14	225	173
MANHUACU	13	161	121
SABARA	13	127	95
UBERABA	12	275	245
FRANCISCO SA	11	117	80
BETIM	10	475	388
JUIZ DE FORA	10	120	93
NOVA SERRANA	9	108	74
SANTA BARBARA DO LESTE	9	21	17

Tabela 2.12: *Top 10 Rodovias Federais (BRs) por Total de Ocorrências (2020).*

Rodovia (BR)	Total de Ocorrências
381	1948
040	1282
116	851
262	793
365	483
050	421
251	201
153	173
267	121
459	70

Tabela 2.13: *Top 10 Pontos Críticos (BR-KM) por Número de Ocorrências (2020).*

Ponto Crítico	Ocorrências
BR-040 KM 511	17
BR-040 KM 510	16
BR-381 KM 490	14
BR-040 KM 517	13
BR-040 KM 522	13
BR-040 KM 524	12
BR-040 KM 532	12
BR-050 KM 69	12
BR-116 KM 415	12
BR-381 KM 483	12

3

Conclusão

Este trabalho analisou os acidentes rodoviários em Minas Gerais durante o ano de 2020. A análise revelou que as rodovias BR-381 e BR-040 são os principais focos de ocorrências, com pontos específicos na BR-040 apresentando uma concentração de acidentes.

Temporalmente, os finais de semana e os horários de pico (manhã e final de tarde) registram o maior volume de colisões. Contudo, é na madrugada que o risco de morte se torna mais elevado. Fatores como pista simples e condições de neblina também foram associados a uma maior severidade dos acidentes, enquanto a "saída de leito carroçável" se destacou como o tipo de ocorrência mais frequente.

Os resultados apontam para a necessidade de intervenções estratégicas. A fiscalização deve ser intensificada nos trechos mais críticos identificados. Campanhas de segurança devem alertar para os perigos específicos da condução na madrugada e em condições climáticas adversas. A longo prazo, o investimento na melhoria da infraestrutura.

Note

Embora limitado ao ano atípico de 2020, os dados se mostraram muito relevantes e completos, com diversas tendências ao longo do EDA, como texto prático gostei muito de realizar o trabalho, agradeço a EstatMG pela oportunidade e quem sabe fazer parte do time!

Obrigado pela Atenção.

